©*школа Опойцева* 2015

Удобства — это, как правило,
вопрос жизни и смерти
математической дисциплины.

ЛА.3. *ЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ И МАТРИЦЫ*

Линейная функция

$$w = \boldsymbol{a} \cdot \boldsymbol{x} = a_1 x_1 + \cdots + a_n x_n$$

принимает постоянные значения $\boldsymbol{a} \cdot \boldsymbol{x} = \beta$ на «плоскостях» параллельных плоскости $\boldsymbol{a} \cdot \boldsymbol{x} = 0$. Действительно, из $\boldsymbol{a} \cdot \boldsymbol{u} = \beta$, $\boldsymbol{a} \cdot \boldsymbol{v} = \beta$ следует $\boldsymbol{a} \cdot (\boldsymbol{u} - \boldsymbol{v}) = 0$, т. е. вектор $\boldsymbol{u} - \boldsymbol{v}$ параллелен (лежит в) плоскости $\boldsymbol{a} \cdot \boldsymbol{x} = 0$.

Множества $\boldsymbol{a} \cdot \boldsymbol{x} = \beta$ называют *гиперплоскостями*[1], чтобы подчеркнуть, что они не являются линейными пространствами, которым положено вместе с векторами содержать их сумму. Другими словами, *гиперплоскость — это плоскость, не проходящая через начало координат.*

Важную роль в анализе играет понятие *линейного оператора*, сопоставляющего вектору $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ вектор $\boldsymbol{y} = \{y_1, \ldots, y_n\}$:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n. \end{cases} \tag{1}$$

Таблицу коэффициентов $A = [a_{ij}]$,

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

[1]Иногда — просто плоскостями, допуская вольность терминологии.

называют *матрицей*, и коротко пишут $\boldsymbol{y} = A\boldsymbol{x}$, определяя тем самым умножение матрицы на вектор. При этом в $\boldsymbol{y} = A\boldsymbol{x}$ под $\boldsymbol{x}$, $\boldsymbol{y}$ подразумеваются вектор-столбцы[2].

Первый индекс у a_{ij} обозначает номер строки, второй — столбца. Совокупность элементов $a_{11}, \ldots, a_{nn}$ называют ***главной диагональю*** матрицы.

Через $\boldsymbol{a_{i:}}$ будем обозначать i-ю строку матрицы A, через $\boldsymbol{a_{:j}}$ — j-й столбец, т. е.

$$\boldsymbol{a_{i:}} = [a_{i1}, \ldots, a_{in}], \qquad \boldsymbol{a_{:j}} = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix}.$$

Двоеточие таким образом заменяет весь список индексов.

Особую роль в линейной алгебре играет понятие *невырожденности*. Матрица называется ***невырожденной***, если её вектор-строки линейно независимы[3].

Умножение матрицы на скаляр γ и сложение матриц $A = [a_{ij}]$ и $B = [b_{ij}]$ определяются так:

$$\gamma A = [\gamma a_{ij}], \qquad A + B = [a_{ij} + b_{ij}].$$

При этом, очевидно,

$$A(\lambda\boldsymbol{x} + \mu\boldsymbol{y}) = \lambda A\boldsymbol{x} + \mu A\boldsymbol{y}.$$

Правило умножения матриц возникает автоматически, если на матрицы смотреть как на линейные операторы. Вектор $\boldsymbol{x}$ под действием оператора B переходит в $\boldsymbol{z} = B\boldsymbol{x}$,

[2]О причинах см. далее.

[3]Забегая вперед, отметим, что из линейной независимости вектор-строк вытекает линейная независимость вектор-столбцов.

а вектор $\boldsymbol{z}$ под действием оператора A переходит в $\boldsymbol{y} = A\boldsymbol{z}$. Результирующее преобразование $\boldsymbol{x}$ в $\boldsymbol{y}$ определяется матрицей $C = AB$ с элементами c_{ij}, формула вычисления которых определяется обыкновенным приведением подобных,

$$y_i = \sum_k a_{ik} \sum_j b_{kj} x_j = \sum_j \sum_k a_{ik} b_{kj} x_j = \sum_j c_{ij} x_j.$$

Следовательно, $C = AB$ есть матрица с элементами

$$\boxed{c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}}\,. \tag{2}$$

Иными словами, c_{ij} равно скалярному произведению i-й вектор-строки $\boldsymbol{a_{i:}}$ на j-й вектор-столбец $\boldsymbol{b}_{\boldsymbol{:j}}$, т. е.

$$\boxed{c_{ij} = \boldsymbol{a_{i:}} \cdot \boldsymbol{b}_{\boldsymbol{:j}}}\,.$$

Полезно также представление

$$\boxed{\boldsymbol{c}_{:j} = A \cdot \boldsymbol{b}_{\boldsymbol{:j}}}\,, \tag{3}$$

показывающее, что j-й столбец матрицы $C = AB$ получается умножением j-го столбца матрицы B на матрицу A. Соотношение (3) означает

$$\begin{bmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{bmatrix} = \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix} b_{1j} + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{bmatrix} b_{nj},$$

т. е. каждый столбец матрицы C представляет собой линейную комбинацию столбцов A. Все это, конечно, тавтология. Но такая перезапись (2) меняет форму восприятия и помогает мыслить более объёмно.

Произведение матриц в общем случае не коммутативно[4], $AB \neq BA$, но ассоциативно, $A(BC) = (AB)C$, и дистрибутивно, $A(B + C) = AB + AC$, что легко проверяется.

Матрица A' с элементами $a'_{ij} = a_{ji}$ называется *транспонированной* к A, и получается как бы пространственным поворотом A вокруг главной диагонали на $180°$, в результате чего столбцы и строки меняются местами. Матрицу называют *симметрической*, если $A = A'$, и — *кососимметрической*, если $A = -A'$.

Для транспонированной матрицы широко используется также обозначение A^T. В общем случае, когда элементы A могут быть комплексными, через A^* обозначают сопряжённую матрицу, которая получается из A транспонированием с последующим комплексным сопряжением всех элементов a_{ij}. Например,

$$\begin{bmatrix} i & 3 \\ 2-i & 1+i \end{bmatrix}^* = \begin{bmatrix} -i & 2+i \\ 3 & 1-i \end{bmatrix}.$$

Таким образом, ***в случае матриц с вещественными элементами****, каковые пока имеются в виду, сопряжённая матрица A^* — есть транспонированная.*

При транспонировании (сопряжении) произведения:

$$(AB)^* = B^* A^*. \tag{4}$$

◀ Доказательство укладывается в одну строчку,

$$(AB)^* = \left[\sum_k a_{ik} b_{kj}\right]^* = \left[\sum_k a_{jk} b_{ki}\right] = \left[\sum_k b^*_{ik} a^*_{kj}\right] = B^* A^*. \quad ▶$$

Обратной к A называют матрицу A^{-1} такую, что

$$A^{-1}A = AA^{-1} = I,$$

[4]В случае $\boxed{AB = BA}$ говорят, что матрицы A и B коммутируют.

где I — *единичная матрица*[5],

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix},$$

определяющая тождественное преобразование $I\boldsymbol{x} \equiv \boldsymbol{x}$. Очевидно, $IA = AI = A$ для любой матрицы A.

Вопрос существования обратной матрицы рассматривается далее. Пока речь идет об определении, но уже из определения вытекают некоторые полезные следствия. В частности, формула аналогичная (4),

$$(AB)^{-1} = B^{-1}A^{-1}, \tag{5}$$

при условии, что матрицы A и B обратимы (невырожденны).

◀ После умножения (5) слева на B, получается $B(AB)^{-1} = A^{-1}$, что после умножения слева на A переходит в очевидное $AB(AB)^{-1} = I$. Разумеется, эта цепочка обладает доказательностью в обратном направлении. Очевидное $AB(AB)^{-1} = I$ после умножения слева на A^{-1}, потом на B^{-1}, — даёт (5)[6]. ▶

Транспонируя (сопрягая) $A^{-1}A = I$, имеем $A^*(A^{-1})^* = I$, откуда

$$(A^*)^{-1} = (A^{-1})^*,$$

что означает *перестановочность* операций транспонирования (сопряжения) и обращения матрицы.

---

[5]Иногда используют обозначение $I = \mathrm{diag}\{1, \dots, 1\}$.

[6]Обычно мы стараемся не засорять текст подобными уточнениями. Концентрация на идейной стороне дела требует жертв.

Прямоугольные матрицы

Матрицу A, имеющую m строк и n столбцов, называют $m \times n$ матрицей. При переходе от квадратных матриц к прямоугольным кое-что теряет смысл, но многое — сохраняется. С некоторыми уточнениями, конечно. Например, произведение $C = AB$ вычисляется по той же формуле (2), но теперь при условии, что *число столбцов A равно числу строк B*. В противном случае AB не имеет смысла.

Понятие прямоугольной матрицы позволяет считать векторы тоже матрицами. При этом надо следить, какую «матрицу» представляет собой вектор, вектор-столбец или вектор-строку. Чтобы матрицу A умножать на $\boldsymbol{x}$ *справа*, надо, чтобы $\boldsymbol{x}$ был обязательно вектор-столбцом. Если $\boldsymbol{x}$, $\boldsymbol{y}$ вектор-столбцы, то скалярное произведение $(\boldsymbol{x}, \boldsymbol{y})$ можно записать в виде произведения матриц $\boldsymbol{x}^*\boldsymbol{y}$.

Для наглядности:

$$[x_1, \ldots, x_n] \cdot \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \sum_i x_i y_i, \qquad \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \cdot [y_1, \ldots, y_n] = \begin{bmatrix} x_1 y_1 & \ldots & x_1 y_n \\ \ldots & \ldots & \ldots \\ x_n y_1 & \ldots & x_n y_n \end{bmatrix}$$

При этом важно не увлекаться общностью, чтобы не потерять различие между векторами и матрицами, которое позволяет считать векторы точками пространства, а матрицы — линейными операторами. Возможность подобной интерпретации составляет главную силу линейной алгебры.